

09/554933

PCT/JP98/05238

日本国特許庁

PATENT OFFICE
JAPANESE GOVERNMENT

20.11.98

セ. A. S. U.

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application:

1997年11月25日

REC'D 15 JAN 1999

PCT

出願番号
Application Number:

平成 9年特許願第323129号

出願人
Applicant(s):

財団法人相模中央化学研究所
株式会社プロテジーン

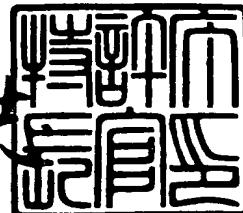
PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED
COMPLIANCE WITH RULE 17.1(a) OR (b)

1998年12月25日

特許庁長官
Commissioner,
Patent Office

伴佐山達志



出証番号 出証特平10-3101942

【書類名】 特許願
【整理番号】 S018128
【提出日】 平成 9年11月25日
【あて先】 特許庁長官殿
【発明の名称】 膜貫通ドメインを有するヒト蛋白質並びにそれをコードするDNA

【請求項の数】 6

【発明者】

【住所又は居所】 神奈川県相模原市若松3-4 6-50
【氏名】 加藤 誠志

【発明者】

【住所又は居所】 東京都葛飾区高砂5-13-11
【氏名】 山口 知子

【発明者】

【住所又は居所】 神奈川県相模原市西大沼4-4-1
【氏名】 関根 伸吾

【特許出願人】

【代表出願人】

【識別番号】 000173762

【郵便番号】 229

【住所又は居所】 神奈川県相模原市西大沼4丁目4番1号

【氏名又は名称】 財団法人相模中央化学研究所

【代表者】 近藤 聖

【電話番号】 0427(42)4791

【特許出願人】

【識別番号】 596134998

【郵便番号】 153

【住所又は居所】 東京都目黒区中町2丁目20番3号

【氏名又は名称】 株式会社プロテジーン

【代表者】 棚井 文雄

【電話番号】 03(3792)1019

【手数料の表示】

【予納台帳番号】 011501

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 膜貫通ドメインを有するヒト蛋白質並びにそれをコードするDNA

【特許請求の範囲】

【請求項 1】 配列番号 1 から配列番号 3 で表されるアミノ酸配列のいずれかを含む蛋白質。

【請求項 2】 請求項 1 記載の蛋白質のいずれかをコードするDNA。

【請求項 3】 配列番号 4 から配列番号 6 で表される塩基配列のいずれかを含むcDNA。

【請求項 4】 配列番号 7 から配列番号 9 で表される塩基配列のいずれかからなる、請求項 3 記載のcDNA。

【請求項 5】 請求項 2 から請求項 4 記載のいずれかのDNAをインビトロ翻訳あるいは真核細胞内で発現しうる発現ベクター。

【請求項 6】 請求項 2 から請求項 4 記載のいずれかのDNAを発現し、請求項 1 記載の蛋白質を生産しうる形質転換真核細胞。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、膜貫通ドメインを有するヒト蛋白質、それをコードしているcDNA、該cDNAの発現ベクター、および該cDNAを発現させた真核細胞に関する。本発明の蛋白質は、医薬品として、あるいは該蛋白質に対する抗体を作製するための抗原として用いることができる。本発明のヒトcDNAは、遺伝子診断用プローブや遺伝子治療用遺伝子源として用いることができる。また、該cDNAがコードしている蛋白質を大量生産するための遺伝子源として用いることができる。これら膜蛋白質遺伝子を導入して膜蛋白質を大量発現させた細胞は、対応するリガンドの検出、新しい低分子医薬のスクリーニングなどに利用できる。

【0002】

【従来の技術】

膜蛋白質は、シグナルレセプター、イオンチャンネル、トランスポーターなど

として、細胞膜を介する物質輸送や情報伝達において重要な役割を担っている。例えば、各種サイトカインに対するレセプター、ナトリウムイオン・カリウムイオン・塩素イオン等に対するイオンチャンネル、糖・アミノ酸等に対するトランスポーターなどが知られており、その多くはすでに遺伝子もクローニングされている。

【0003】

これらの膜蛋白質の異常は、これまで原因不明であった多くの病気と関連していることがわかつてきた。例えば、囊胞性線維症の原因遺伝子として12個の膜貫通ドメインを有する膜蛋白質の遺伝子が同定された [Rommens, J. M. et al., Science 245: 1059-1065 (1989)]。また、いくつかの膜蛋白質は、ウイルスが細胞に感染する際のレセプターとして働いていることがわかつてきた。、例えば、HIV-1は、T細胞膜上の膜蛋白質、CD4抗原と7個の膜貫通ドメインを有する膜蛋白質ヒュージンを介して細胞内に感染することが示された [Feng, Y. et al., Science 272: 872-877 (1996)]。従って、新しい膜蛋白質が見い出せれば、多くの病気の原因解明につながるものと期待され、膜蛋白質をコードする新たな遺伝子の単離が望まれている。

【0004】

従来、膜蛋白質は、精製することが困難なので、遺伝子の方からのアプローチによって単離されたものが多い。一般的な方法は、cDNAライブラリーを真核細胞に導入して、cDNAを発現させたのち、目的とする膜蛋白質を膜上に発現している細胞を、抗体を用いる免疫学的な手法や膜の透過性の変化を生理学的な手法で検出する、いわゆる発現クローニングである。しかしこの方法では機能のわかつた膜蛋白質の遺伝子しかクローニングできない。

【0005】

一般に膜蛋白質は、蛋白質内部に疎水性の膜貫通ドメインを有しており、リボソームで合成された後、このドメインがリン脂質膜内に留まり膜にトラップされる。従って、完全長cDNAの全塩基配列を決定してやり、そのcDNAがコードしている蛋白質のアミノ酸配列の中に疎水性の高い膜貫通ドメインが存在すれ

ば、そのcDNAは膜蛋白質をコードしていると考えられる。

【0006】

【発明が解決しようとする課題】

本発明の目的は、膜貫通ドメインを有する新規のヒト蛋白質、該蛋白質をコードするDNA、該cDNAの発現ベクター、および該cDNAを発現しうる形質転換真核細胞を提供することである。

【0007】

【課題を解決するための手段】

本発明者らは鋭意研究の結果、ヒト完全長cDNAバンクの中から膜貫通ドメインを有する蛋白質をコードするcDNAをクローニングし、本発明を完成した。すなわち、本発明は膜貫通ドメインを有するヒト蛋白質である、配列番号1から配列番号3で表されるアミノ酸配列のいずれかを含む蛋白質を提供する。また本発明は上記蛋白質をコードするDNA、例えば配列番号4から配列番号9で表される塩基配列のいずれかを含むcDNA並びに該cDNAを発現しうる形質転換真核細胞を提供する。

【0008】

【発明の実施の形態】

本発明の蛋白質は、ヒトの臓器、細胞株などから単離する方法、本発明のアミノ酸配列に基づき化学合成によってペプチドを調製する方法、あるいは本発明の膜貫通ドメインをコードするDNAを用いて組換えDNA技術で生産する方法などにより取得することができるが、組換えDNA技術で取得する方法が好ましく用いられる。例えば、本発明のcDNAを有するベクターからインビトロ転写によってRNAを調製し、これを鋳型としてインビトロ翻訳を行なうことによりインビトロで蛋白質を発現できる。また翻訳領域を公知の方法により適当な発現ベクターに組換えてやれば、大腸菌、枯草菌等の原核細胞や、酵母、昆虫細胞、哺乳動物細胞等の真核細胞で、コードしている蛋白質を大量に発現させることができる。

【0009】

本発明の蛋白質を、インビトロ翻訳でDNAを発現させて生産させる場合には

、該cDNAの翻訳領域を、RNAポリメラーゼプロモーターを有するベクターに組換え、プロモーターに対応するRNAポリメラーゼを含む、ウサギ網状赤血球溶解物や小麦胚芽抽出物などのインビトロ翻訳系に添加してやれば、本発明の蛋白質をインビトロで生産することができる。RNAポリメラーゼプロモーターとしては、T7、T3、SP6などが例示できる。これらのRNAポリメラーゼプロモーターを含むベクターとしては、pKA1、pCDM8、pT3/T7 18、pT7/3 19、pBluescript IIなどが例示できる。また、反応系にイヌ臍臍ミクロソームなどを添加してやれば、本発明の膜蛋白質をミクロソーム膜に組み込まれた形で発現することができる。

【0010】

本発明の蛋白質を、大腸菌などの微生物でDNAを発現させて生産させる場合には、微生物中で複製可能なオリジン、プロモーター、リボソーム結合部位、cDNAクローニング部位、ターミネーター等を有する発現ベクターに、本発明のcDNAの翻訳領域を組換えた発現ベクターを作成し、該発現ベクターで宿主細胞を形質転換したのち、得られた形質転換体を培養してやれば、該cDNAがコードしている蛋白質を微生物内で大量生産することができる。この際、任意の翻訳領域の前後に開始コドンと停止コドンを付加して発現させてやれば、任意の領域を含む蛋白質断片を得ることができる。あるいは、他の蛋白質との融合蛋白質として発現させることもできる。該融合蛋白質を適当なプロテアーゼで切断することによって該cDNAがコードする蛋白質部分のみを取得することもできる。大腸菌用発現ベクターとしては、pUC系、pBluescript II、pET発現システム、pGEX発現システムなどが例示できる。

【0011】

本発明の蛋白質を、真核細胞でDNAを発現させて生産させる場合には、該cDNAの翻訳領域を、プロモーター、スプライシング領域、ポリ(A)付加部位等を有する真核細胞用発現ベクターに組換え、真核細胞内に導入してやれば、本発明の蛋白質を膜蛋白質として細胞膜表面上で生産することができる。発現ベクターとしては、pKA1、pCDM8、pSVK3、pMSG、pSVL、pBK-CMV、pBK-RSV、EBVベクター、pRS、pYES2などが例示

できる。真核細胞としては、サル腎臓細胞COS7、チャイニーズハムスター卵巣細胞CHOなどの哺乳動物培養細胞、出芽酵母、分裂酵母、カイコ細胞、アフリカツメガエル卵細胞などが一般に用いられるが、本蛋白質を膜表面に発現できるものであれば、いかなる真核細胞でもよい。発現ベクターを真核細胞に導入するには、電気穿孔法、リン酸カルシウム法、リポソーム法、DEAEデキストラン法など公知の方法を用いることができる。

【0012】

本発明の蛋白質を原核細胞や真核細胞で発現させたのち、培養物から目的蛋白質を単離精製するためには、公知の分離操作を組み合わせて行うことができる。例えば、尿素などの変性剤や界面活性剤による処理、超音波処理、酵素消化、塩析や溶媒沈殿法、透析、遠心分離、限外濾過、ゲル濾過、SDS-PAGE、等電点電気泳動、イオン交換クロマトグラフィー、疎水性クロマトグラフィー、アフィニティーコロマトグラフィー、逆相クロマトグラフィーなどがあげられる。

【0013】

本発明の蛋白質には、配列番号1から配列番号3で表されるアミノ酸配列のいかなる部分アミノ酸配列を含むペプチド断片（5アミノ酸残基以上）も含まれる。これらのペプチド断片は抗体を作製するための抗原として用いることができる。また、本発明の蛋白質の中でシグナル配列を有するものは、シグナル配列が除去された後、成熟蛋白質の形で細胞表面に出てくる。したがって、これらの成熟蛋白質は本発明の蛋白質の範疇にはいる。成熟蛋白質のN末端アミノ酸配列は、シグナル配列切断部位決定法〔特開平8-187100〕を用いて容易に求めることができる。また、いくつかの膜蛋白質は、細胞表面でプロセシングを受けて分泌型となる。このような分泌型となった蛋白質あるいはペプチドも本発明の蛋白質の範疇にはいる。アミノ酸配列の中に糖鎖結合部位が存在すると、適当な真核細胞で発現させれば糖鎖が付加した蛋白質が得られる。したがって、このような糖鎖が付加した蛋白質あるいはペプチドも本発明の蛋白質の範疇にはいる。

【0014】

本発明のDNAには、上記蛋白質をコードするすべてのDNAが含まれる。該DNAは、化学合成による方法、cDNAクローニングによる方法などを用いて

取得することができる。

【0015】

本発明のcDNAは、例えばヒト細胞由来cDNAライブラリーからクローン化することができる。cDNAはヒト細胞から抽出したポリ(A)⁺RNAを鑄型として合成する。ヒト細胞としては、人体から手術などによって摘出されたものでも培養細胞でも良い。cDNAは、岡山-Berg法 [Okayama, H. and Berg, P., Mol. Cell. Biol. 2:161-170 (1982)]、Gubler-Hoffmann法 [Gubler, U. and Hoffmann, J., Gene 25:263-269 (1983)]などいかなる方法を用いて合成してもよいが、完全長クローンを効率的に得るためには、実施例にあげたようなキャッシング法 [Kato, S. et al., Gene 163:193-196 (1995)]を用いることが望ましい。また市販のヒトcDNAライブラリーを用いることもできる。cDNAライブラリーから本発明のcDNAをクローン化するには、本発明のcDNAの任意の部分の塩基配列に基づいてオリゴヌクレオチドを合成し、これをプローブとして用いて、公知の方法によりコロニーあるいはプラークハイブリダイゼーションによるスクリーニングを行えばよい。また、目的とするcDNA断片の両末端にハイブリダイズするオリゴヌクレオチドを合成し、これをプライマーとして用いて、ヒト細胞から単離したmRNAからRT-PCR法により、本発明のcDNA断片を調製することもできる。

【0016】

本発明のcDNAは、配列番号4から配列番号6で表される塩基配列あるいは配列番号7から配列番号9で表される塩基配列のいずれかを含むことを特徴とするものである。それぞれのクローン番号(HP番号)、cDNAクローンが得られた細胞、cDNAの全塩基数、コードしている蛋白質のアミノ酸残基数をそれぞれ表1にまとめて示した。

【0017】

【表1】

表1

配列番号	HP番号	細胞	塩基数	アミノ酸 残基数
1、4、7	HP01207	胃癌	2938	269
2、5、8	HP01862	胃癌	2290	311
3、6、9	HP10493	PMA-U937	3705	383

【0018】

なお、配列番号4から配列番号9のいずれかに記載のcDNAの塩基配列に基づいて合成したオリゴヌクレオチドプローブを用いて、本発明で用いたヒト細胞株やヒト組織から作製したcDNAライブラリーをスクリーニングすることにより、本発明のcDNAと同一のクローンを容易に得ることができる。

【0019】

一般にヒト遺伝子は個体差による多型が頻繁に認められる。従って配列番号4から配列番号9において、1又は複数個のヌクレオチドの付加、欠失および／又は他のヌクレオチドによる置換がなされているcDNAも本発明の範疇にはいる。

【0020】

同様に、これらの変更によって生じる、1又は複数個のアミノ酸の付加、欠失および／又は他のアミノ酸による置換がなされている蛋白質も、配列番号1から配列番号3で表されるアミノ酸配列を有するそれぞれの蛋白質の活性を有する限り、本発明の範疇に入る。

【0021】

本発明のcDNAには、配列番号4から配列番号6で表される塩基配列あるいは配列番号7から配列番号9で表される塩基配列のいかなる部分塩基配列を含むcDNA断片(10bp以上)も含まれる。また、センス鎖およびアンチセンス鎖からなるDNA断片もこの範疇にはいる。これらのDNA断片は遺伝子診断用のプローブとして用いることができる。

【0022】

【実施例】

次に実施例により発明を具体的に説明するが、本発明はこれらの例に限定されるものではない。DNAの組換えに関する基本的な操作および酵素反応は、文献["Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Laboratory, 1989]に従った。制限酵素および各種修飾酵素は特に記載の無い場合宝酒造社製のものを用いた。各酵素反応の緩衝液組成、並びに反応条件は付属の説明書に従った。cDNA合成は文献[Kato, S. et al., Gene 150: 243-250 (1994)]に従った。

【0023】

(1) ポリ(A)⁺RNAの調製

mRNAを抽出するためのヒト細胞として、ホルボールエステルで刺激した組織球リンホーマ細胞株U937 (ATCC CRL 1593)、手術によって摘出された胃癌組織を用いた。細胞株の培養は、常法に従って行った。

【0024】

ヒト細胞約1gを5.5Mグアニジウムチオシアネート溶液20ml中でホモジナイズした後、文献[Okayama, H. et al., "Methods in Enzymology" Vol. 164, Academic Press, 1987]に従い、総mRNAを調製した。これを20mMトリス塩酸緩衝液(pH7.6)、0.5MNaCl、1mMEDTAで洗浄したオリゴdTセルロースカラムにかけ、上掲文献に従いポリ(A)⁺RNAを得た。

【0025】

(2) cDNAライブラリーの作製

上記ポリ(A)⁺RNA 10μgを100mMトリス塩酸緩衝液(pH8)に溶解し、RNaseを含まないバクテリア由来アルカリホスファターゼ1単位を添加し、37℃1時間反応させた。反応液をフェノール抽出後、エタノール沈殿を行ない、ペレットを50mM酢酸ナトリウム(pH6)、1mMEDTA、0.1%2-メルカプトエタノール、0.01%Triton X-100溶液

に溶解した。これに、タバコ由来酸ピロホスファターゼ（エピセンターテクノロジーズ社製）1単位を添加して、総量100μlで37℃1時間反応させた。反応液をフェノール抽出後、エタノール沈殿を行ない、ペレットを水に溶解し、脱キップ処理したポリ(A)⁺RNA溶液を得た。

【0026】

脱キップ処理したポリ(A)⁺RNA、DNA-RNAキメラオリゴヌクレオチド(5'-dG-dG-dG-dG-dA-dA-dT-dT-dC-dG-dA-G-G-A-3')3nmolを50mMトリス塩酸緩衝液(pH7.5)、0.5mMATP、5mMMgCl₂、10mM2-メルカプトエタノール、25%ポリエチレングリコール水溶液に溶解し、T4 RNAリガーゼ50単位を添加し、総量30μlで20℃12時間反応させた。反応液をフェノール抽出後、エタノール沈殿を行ない、ペレットを水に溶解し、キメラオリゴキップ付加ポリ(A)⁺RNAを得た。

【0027】

本発明者らが開発したベクターpKA1(特開平4-117292号公報)をKpnIで消化後、末端転移酵素により約60個のdTテールを付加した。これをEcoRV消化して片側のdTテールを除去したものをベクタープライマーとして用いた。

【0028】

先に調製したキメラオリゴキップ付加ポリ(A)⁺RNA 6μgを、ベクタープライマー1.2μgとアニールさせた後、50mMトリス塩酸緩衝液(pH8.3)、75mMKCl、3mMMgCl₂、10mMジチオスレイトール、1.25mMDNTP(dATP+dCTP+dGTP+dTTP)溶液に溶解し、逆転写酵素(GIBCO-BRL社製)200単位を添加し、総量20μlで42℃1時間反応させた。反応液をフェノール抽出後、エタノール沈殿を行ない、ペレットを50mMトリス塩酸緩衝液(pH7.5)、100mMNaCl、10mMMgCl₂、1mMジチオスレイトール溶液に溶解した。これにEcoRI 100単位を添加し、総量20μlで37℃1時間反応させた。反応液をフェノール抽出後、エタノール沈殿を行ない、ペレットを20mMトリス塩酸緩

衝液 (pH 7.5)、100 mM KCl、4 mM MgCl₂、10 mM (NH₄)₂SO₄、50 μg/ml 牛血清アルブミン溶液に溶解した。これに大腸菌DNAリガーゼ60単位を添加し、16°C 16時間反応させた。反応液に 2 mM dNTP 2 μl、大腸菌DNAポリメラーゼI 4単位、大腸菌RNase H 0.1単位を添加し、12°C 1時間ついで 22°C 1時間反応させた。

【0029】

次いでcDNA合成反応液を用いて大腸菌DH12S (GIBCO-BRL社製) の形質転換を行なった。形質転換はエレクトロポレーション法によって行なった。形質転換体の一部を 100 μg/ml アンピシリン含有 2×YT 寒天培地上に蒔いて 37°C 一晩培養した。寒天上に生じた任意のコロニーを拾い 100 μg/ml アンピシリン含有 2×YT 培地 2 ml に接種して 37°C で一晩培養した。培養液を遠心して、菌体からアルカリリシス法によりプラスミドDNAを調製した。プラスミドDNAは EcoRI と NotI で二重消化した後、0.8% アガロースゲル電気泳動を行ない cDNA インサートの大きさを求めた。また、得られたプラスミドを鋳型にして、蛍光色素で標識した M13 ユニバーサルプライマーと Taq ポリメラーゼ (アプライドバイオシステムズ社製キット) を用いてシーケンス反応を行なった後、蛍光DNAシーケンサー (アプライドバイオシステムズ社) にかけて cDNA の 5' 末端約 400 bp の塩基配列を決定した。配列データはホモ・プロテイン cDNA バンクデータベースとしてファイル化した。

【0030】

(3) 膜貫通ドメインを有する蛋白質をコードしている cDNA の選択

ホモ・プロテイン cDNA バンクに登録された塩基配列を 3 フレームのアミノ酸配列に変換し、開始コドンから始まるオープンリーディングフレーム (ORF) の有無を調べた。次いで ORF がコードしている部分の N 末端に分泌蛋白質に特有なシグナル配列が認められるものを選択した。これらのクローンについては、エキソヌクレアーゼ I II による欠失法を用いて、5' 並びに 3' 両方向からシーケンシングを行い、全塩基配列の決定を行った。ORF がコードしている蛋白質について、Kyte-Doolittle の方法 [Kyte, J & Do

o l i t t l e, R. F., J. Mol. Biol. 157: 105-132 (1982)] により、疎水性／親水性プロフィールを求め、疎水性領域の有無を調べた。コードしている蛋白質のアミノ酸配列中に膜貫通ドメインと思われる疎水的な領域がある場合には、この蛋白質は膜蛋白質であると見なした。

【0031】

(4) 分泌シグナル配列あるいは膜貫通ドメインの機能確認

上記工程の結果得られた分泌蛋白質候補クローンについて、N末端の疎水性領域が分泌シグナル配列として機能することを、文献記載の方法 [Y o k o y a m a-K o b a y a s h i, M. et al., Gene 163: 193-196 (1995)] によって確認した。まずターゲット c D N A を含んでいるプラスミドを、分泌シグナル配列をコードしていると考えられる部分の下流に存在する適当な制限酵素部位で切断した。もしこの制限酵素部位が突出末端である場合には、クレノウ処理やマングビーンヌクレアーゼ処理によって平滑末端にした。さらに H i n d I I I による消化を行い、S V 4 0 プロモーターとその下流に分泌シグナル配列をコードしている c D N A を含むDNA断片をアガロースゲル電気泳動によって単離した。この断片を、p S S D 3 (D D B J / E M B L / G e n B a n k 登録番号AB007632) の H i n d I I I と、ウロキナーゼのコーディングフレームと合うように選択した制限酵素部位の間に挿入し、ターゲット c D N A の分泌シグナル配列部分とウロキナーゼプロテアーゼドメインの融合蛋白質を発現するためのベクターを構築した。

【0032】

融合蛋白質発現ベクターを有する大腸菌（宿主：J M 1 0 9）を 1 0 0 μ g / m l アンピシリン含有 2 × Y T 培地 2 m l 中で 3 7 °C 2 時間培養した後、ヘルパー・ファージM 1 3 K O 7 (50 μ l) を添加し、3 7 °C で一晩培養した。遠心によって分離した上澄からポリエチレングリコール沈殿によって一本鎖ファージ粒子を得た。これを 1 0 0 μ l の 1 mM トリス - 0. 1 mM E D T A 、p H 8 (T E) に懸濁した。また対照として、p S S D 3 、並びにウロキナーゼの完全長 c D N A を含むベクター p K A 1 - U P A [Y o k o y a m a - K o b a y a s h i, M. et al., Gene 163: 193-196 (1995)] か

ら同様にして調製した一本鎖ファージ粒子懸濁液を用いた。

【0033】

サル腎臓由来培養細胞COS7は、10%ウシ胎児血清を含むダルベッコ改変イーグル(DMEM)培地中、5%CO₂存在下、37℃で培養した。1×10⁵個のCOS7細胞を6穴プレート(ヌンク社、穴の直径3cm)に植え、5%CO₂存在下、37℃で22時間培養した。培地除去後、リン酸緩衝液で細胞表面を洗浄し、さらに50mMトリス塩酸(pH7.5)を含むDMEM(TDMEM)で再度洗浄した。この細胞に一本鎖ファージ懸濁液1μl、DMEM培地0.6ml、TRANSFECTAMTM(IBF社)3μlを懸濁したものを添加し、5%CO₂存在下、37℃で3時間培養した。サンプル液を除去後、TDMEMで細胞表面を洗浄し、10%ウシ胎児血清含有DMEMを1穴あたり2ml加え、5%CO₂存在下、37℃にて2日間培養した。

【0034】

2%ウシフィブリノーゲン(マイルス社)、0.5%アガロース、1mM塩化カルシウムを含む50mMリン酸緩衝液(pH7.4)10mlに10単位のヒトトロンビン(持田製薬)を加え、直径9cmのプレート中で固化させ、フィブリンプレートを調製した。トランスフェクションしたCOS7細胞の培養上清10μlをフィブリンプレートに載せ、37℃15時間インキュベートした。フィブリンプレート上に溶解円が現れたら、cDNA断片が分泌シグナル配列として機能するアミノ酸配列をコードしていることを意味する。一方、溶解円を形成しない場合には、細胞を十分洗浄した後、フィブリンシートを細胞の上に乗せて、37℃15時間インキュベートした。もし、フィブリンシートに溶解部分が生じたら、細胞表面にウロキナーゼ活性が発現したことを示す。すなわち、cDNA断片は、膜貫通ドメインをコードしていることを意味する。

【0035】

(5) インビトロ翻訳による蛋白質合成

本発明のcDNAを有するプラスミドベクターを用いて、T_NTウサギ網状赤血球溶解物キット(プロメガ社製)によるインビトロ転写/翻訳を行なった。この際 [³⁵S]メチオニンを添加し、発現産物をラジオアイソトープでラベルした

。いずれの反応もキットに付属のプロトコールに従って行なった。プラスミド2 μ gを、T_{NT}ウサギ網状赤血球溶解物12.5 μ l、緩衝液（キットに付属）0.5 μ l、アミノ酸混合液（メチオニンを含まない）2 μ l、[³⁵S]メチオニン（アマーシャム社）2 μ l（0.37MBq/ μ l）、T7RNAポリメラーゼ0.5 μ l、RNasin 20Uを含む総量25 μ lの反応液中で30°Cで90分間反応させた。反応液3 μ lにSDSサンプリングバッファー（125mMトリス塩酸緩衝液、pH6.8、120mM2-メルカプトエタノール、2%SDS溶液、0.025%プロモフェノールブルー、20%グリセロール）2 μ lを加え、95°C3分間加熱処理した後、SDS-PAGEにかけた。オートラジオグラフィーを行ない、翻訳産物の分子量を求めた。

【0036】

(6) COS7による発現

本発明の蛋白質の発現ベクターを有する大腸菌に、ヘルパーファージM13KO7を感染させ、上記の方法で一本鎖ファージ粒子を得た。得られたファージを用いて上記の方法によりサル腎臓由来培養細胞COS7に各発現ベクターを導入した。5%CO₂存在下、37°Cで2日間培養したのち、[³⁵S]システインあるいは[³⁵S]メチオニンを含む培地中で1時間培養した。細胞を集め溶解した後、SDS-PAGEにかけたところ、COS7細胞には存在しない、各蛋白質の発現産物に相当するバンドが認められた。

【0037】

(7) クローン例

<HP01207> (配列番号1、4、7)

ヒト胃癌cDNAライブラリーから得られたクローンHP01207のcDNAインサートの全塩基配列を決定したところ、100bpの5'非翻訳領域、810bpのORF、2028bpの3'非翻訳領域からなる構造を有していた。ORFは269アミノ酸残基からなる蛋白質をコードしており、7箇所の推定膜貫通ドメインが存在した。図1にKyte-Doolittleの方法で求めた本蛋白質の疎水性/親水性プロフィールを示す。インビトロ翻訳の結果、高分子量のスマアな翻訳産物が生成した。

【0038】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、マウス S u r f - 4 蛋白質 (P I R アクセション番号 A 3 4 7 2 7) と類似性を有していた。表 2 に、本発明のヒト蛋白質 (H P) とマウス S u r f - 4 蛋白質 (MM) のアミノ酸配列の比較を示す。* は本発明の蛋白質と同一アミノ酸残基を、. は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。両者は、全領域で 99.3 % の相同性を有していた。

【0039】

【表 2】

表 2

HS	MGQN DLMGTAEDFADQFLRVTKQYLPHV AR LCLISTFLEDGIR MWFQWSEQR DYIDTTWN
MM	MGQN DLMGTAEDFADQFLRVTKQYLPHV AR LCLISTFLEDGIR MWFQWSEQR DYIDTTWS
HS	CGYLLASSFVFLNLLGQLTGCVLVLSRN FVQYACFGLFGII ALQTI AYSILWDLKFLMRN
MM	CGYLLASSFVFLNLLGQLTGCVLVLSRN FVQYACFGLFGII ALQTI AYSILWDLKFLMRN
HS	LALGGGLLLL A ESRSEGKSMFAGVPTMRESSPKQYMQ LGGRVLLVLMFMTLLHFDASFF
MM	LALGGGLLLL A ESRSEGKSMFAGVPTMRESSPKQYMQ LGGRVLLVLMFMTLLHFDASFF
HS	SIVQNIVGTALMILVAIGF KTKLAALTLVVWLFAINVYFNAFWTIPVYKPMHDFLKYDFF
MM	SIIQNIVGTALMILVAIGF KTKLAALTLVVWLFAINVYFNAFWTIPVYKPMHDFLKYDFF
HS	QTMSVIGGLLVVALGPGGVSMDEKKKEW
MM	QTMSVIGGLLVVALGPGGVSMDEKKKEW

【0040】

また、本 c D N A の塩基配列を用いて G e n B a n k を検索したところ、1 2

2番目から883番目までの762bpと98.6%の類似性を示す塩基配列（GenBankアクセション番号Y14820）が登録されていたが、本蛋白質の断片をコードするものである。

【0041】

マウスSurf-4蛋白質は、マウスsurf e i t locusにコードされている蛋白質の一つであり、細胞の生存に必須であるハウスキーピング蛋白質と考えられている[Huxley, C. et al., Mol. Cell. Biol. 10: 605-614 (1990)]。

【0042】

<HP01862>（配列番号2、5、8）

ヒト胃癌cDNAライブラリーから得られたクローンHP01862のcDNAインサートの全塩基配列を決定したところ、80bpの5'非翻訳領域、936bpのORF、1274bpの3'非翻訳領域からなる構造を有していた。ORFは311アミノ酸残基からなる蛋白質をコードしており、7箇所の膜貫通ドメインが存在した。図2にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロフィールを示す。インビトロ翻訳の結果、高分子量のスマートな翻訳産物が生成した。

【0043】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、ラットNMDAレセプターグルタミン酸結合サブユニット（Genbankアクセション番号S19586）と類似性を有していた。表3に、本発明のヒト蛋白質（HP）とラットNMDAレセプターグルタミン酸結合サブユニット（RN）のアミノ酸配列の比較を示す。-はギャップを、*は本発明の蛋白質と同一アミノ酸残基を、.は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。両者は、41.0%の相同性を有していた。

【0044】

【表3】

表3

HS

MSNPSAPPYEDRNP

RN MKRVWSLGTAILPQTLAILWGHKPLCLPMFSLPTLGPHTHRPLSSPLPMVNQGIPMVPV

HS LYPGPLPPGGYQGQPSVLPGGYPAYPGYPQPGYGHAGYPQPMPPTHPMMPMNYGPGHGYDG

** * *. * . *. **. *. . . * . * * . * * **.

RN PITRWLPLKDLLKEATHQGHYPQSP-FPPNPYQQPPPQDQPGSPQHGNYQEEGPPSYDN

HS EERA VSDSFGPGEWDDRKV RHTFIRKV YSIISVQLLITV AIIAIFTF VEPVSAF VRRNVA

.. * . * * . . . * . . . * . . . * . . . * . . . * . . . * . . . * . . . * .

RN QD-----FPSVNW-DKSIRQAFIRKVFLVLTQLS VTLSTVAIFTF VGEVKGFVRANVW

HS VYYVSYAVFVVTYLILACCQGP RRFPWNIILLT LFTFAMGFMTGTISSMYQTKAVIIAM

. * . * . * . . . * . . . * . . . * . . . * . . . * . . . * . . . * . . . * .

RN TYYVSYA IFFISLIVLSCCGDFRKHPWNLVALSILTISLSY MVGMIA SFYNT EAVIMAV

HS IITAVVSISVTIFCFQTKV DFTSCTGLFCVLGI VLLVTGIVTSIVLYFQYVYWLHMLYAA

**. . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * .

RN GITTAVCFTVVIFSMQTRYDFTSCMGVLLVSVVVLFIFAIL---CIFIRNRI-LEIVYAS

HS LGAICFTLFLAYDTQLVLGNRKHTISPEDYITGALQIYTDIIYIIFTFVLQLMGDRN

***. ** *** ***. ***. . . ***. *. . ***. *** * . * . * .

RN LGALLFTCFLAVDTQLLLGNKQLSLSPEEYVFAALNLYTDIINIFLYILTIIGRSQGIGQ

【0045】

また、本cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同意を有するもの（例えば、アクセション番号H06014）が存在したが、いずれも本cDNAより短く、開始コドンから含んでいるものは見いだせなかった。

【0046】

ラットNMDAレセプターグルタミン酸結合サブユニットは、脳に特異的に存在するNMDAレセプター複合体のサブユニットの一つとして見いだされた [Kumar, K. N. et al., Nature 354: 70-73 (1991)]。本発明の蛋白質は、チャンネルやトランスポーターに特徴的な7個の膜

貫通ドメインを有することから、チャネルやトランスポーターとしての役割を担っていると思われる。

【0047】

<HP10493> (配列番号3、6、9)

ヒトリンホーマU937 cDNAライブラリーから得られたクローンHP10493のcDNAインサートの全塩基配列を決定したところ、123bpの5'非翻訳領域、1152bpのORF、2430bpの3'非翻訳領域からなる構造を有していた。ORFは383アミノ酸残基からなる蛋白質をコードしており、N末端に1箇所の膜貫通ドメインが存在した。図3にKyte-Doolittle法の方法で求めた本蛋白質の疎水性/親水性プロフィールを示す。本蛋白質のN末端44アミノ酸残基をコードしているcDNA部分を含むHindIII-AccI断片をpSSD3のHindIII-PmaCI部位に挿入した発現ベクターをCOS7細胞に導入したところ、細胞表面にウロキナーゼ活性が認められ、本蛋白質はII型膜蛋白質であることが示された。インビトロ翻訳の結果、ORFから予想される分子量43,001とほぼ同じ43kDaの翻訳産物が生成した。

【0048】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したが、類似性を有する既知蛋白質はなかった。モチーフ配列を検索したところ、175番目のヒスチジンは、トリプシン型のセリンプロテアーゼの活性部位である可能性が高い。従って、本蛋白質は膜型のプロテアーゼの可能性がある。また、本cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同意を有するもの（例えば、アクセス番号R81003）が存在したが、不明瞭な配列が多く、本cDNAと同じORFは見いだせなかった。

【0049】

【発明の効果】

本発明は膜貫通ドメインを有するヒト蛋白質、それをコードしているcDNA、該cDNAの発現ベクター、および該cDNAを発現させた真核細胞を提供する。本発明の蛋白質は、いずれも細胞膜に存在するので、細胞の増殖や分化を制

御している蛋白質と考えられる。したがって、本発明の蛋白質は、細胞の増殖や分化の制御に関わる制癌剤などの医薬品として、あるいは該蛋白質に対する抗体を作製するための抗原として用いることができる。本発明のcDNAは、遺伝子診断用プローブや遺伝子治療用遺伝子源として用いることができる。また、該DNAを用いることにより、該蛋白質を大量に発現することができる。これら膜蛋白質遺伝子を導入して膜蛋白質を大量発現させた細胞は、対応するリガンドの検出、新しい低分子医薬のスクリーニングなどに利用できる。

【0050】

【配列表】

配列番号：1

配列の長さ：269

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル：N○

起源：

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名：HP01207

配列

Met Gly Gln Asn Asp Leu Met Gly Thr Ala Glu Asp Phe Ala Asp Gln

1 5 10 15

Phe Leu Arg Val Thr Lys Gln Tyr Leu Pro His Val Ala Arg Leu Cys

20 25 30

Leu Ile Ser Thr Phe Leu Glu Asp Gly Ile Arg Met Trp Phe Gln Trp

35 40 45

Ser Glu Gln Arg Asp Tyr Ile Asp Thr Thr Trp Asn Cys Gly Tyr Leu

50 55 60

Leu Ala Ser Ser Phe Val Phe Leu Asn Leu Leu Gly Gln Leu Thr Gly

65	70	75	80
Cys Val Leu Val Leu Ser Arg Asn Phe Val Gln Tyr Ala Cys Phe Gly			
85	90	95	
Leu Phe Gly Ile Ile Ala Leu Gln Thr Ile Ala Tyr Ser Ile Leu Trp			
100	105	110	
Asp Leu Lys Phe Leu Met Arg Asn Leu Ala Leu Gly Gly Gly Leu Leu			
115	120	125	
Leu Leu Leu Ala Glu Ser Arg Ser Glu Gly Lys Ser Met Phe Ala Gly			
130	135	140	
Val Pro Thr Met Arg Glu Ser Ser Pro Lys Gln Tyr Met Gln Leu Gly			
145	150	155	160
Gly Arg Val Leu Leu Val Leu Met Phe Met Thr Leu Leu His Phe Asp			
165	170	175	
Ala Ser Phe Phe Ser Ile Val Gln Asn Ile Val Gly Thr Ala Leu Met			
180	185	190	
Ile Leu Val Ala Ile Gly Phe Lys Thr Lys Leu Ala Ala Leu Thr Leu			
195	200	205	
Val Val Trp Leu Phe Ala Ile Asn Val Tyr Phe Asn Ala Phe Trp Thr			
210	215	220	
Ile Pro Val Tyr Lys Pro Met His Asp Phe Leu Lys Tyr Asp Phe Phe			
225	230	235	240
Gln Thr Met Ser Val Ile Gly Gly Leu Leu Leu Val Val Ala Leu Gly			
245	250	255	
Pro Gly Gly Val Ser Met Asp Glu Lys Lys Lys Glu Trp			
260	265		

【0051】

配列番号：2

配列の長さ：311

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル：No

起源：

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名：HP01862

配列

Met	Ser	Asn	Pro	Ser	Ala	Pro	Pro	Pro	Tyr	Glu	Asp	Arg	Asn	Pro	Leu
1															15
Tyr	Pro	Gly	Pro	Leu	Pro	Pro	Gly	Gly	Tyr	Gly	Gln	Pro	Ser	Val	Leu
															30
20															
Pro	Gly	Gly	Tyr	Pro	Ala	Tyr	Pro	Gly	Tyr	Pro	Gln	Pro	Gly	Tyr	Gly
35															45
His	Pro	Ala	Gly	Tyr	Pro	Gln	Pro	Met	Pro	Pro	Thr	His	Pro	Met	Pro
50															
Met	Asn	Tyr	Gly	Pro	Gly	His	Gly	Tyr	Asp	Gly	Glu	Glu	Arg	Ala	Val
65															80
Ser	Asp	Ser	Phe	Gly	Pro	Gly	Glu	Trp	Asp	Asp	Arg	Lys	Val	Arg	His
															95
85															
Thr	Phe	Ile	Arg	Lys	Val	Tyr	Ser	Ile	Ile	Ser	Val	Gln	Leu	Leu	Ile
100															
Thr	Val	Ala	Ile	Ile	Ala	Ile	Phe	Thr	Phe	Val	Glu	Pro	Val	Ser	Ala
115															
Phe	Val	Arg	Arg	Asn	Val	Ala	Val	Tyr	Tyr	Val	Ser	Tyr	Ala	Val	Phe
130															
Val	Val	Thr	Tyr	Leu	Ile	Leu	Ala	Cys	Cys	Gln	Gly	Pro	Arg	Arg	Arg
145															160
Phe	Pro	Trp	Asn	Ile	Ile	Leu	Leu	Thr	Leu	Phe	Thr	Phe	Ala	Met	Gly

165	170	175
Phe Met Thr Gly Thr Ile Ser Ser Met Tyr Gln Thr Lys Ala Val Ile		
180	185	190
Ile Ala Met Ile Ile Thr Ala Val Val Ser Ile Ser Val Thr Ile Phe		
195	200	205
Cys Phe Gln Thr Lys Val Asp Phe Thr Ser Cys Thr Gly Leu Phe Cys		
210	215	220
Val Leu Gly Ile Val Leu Leu Val Thr Gly Ile Val Thr Ser Ile Val		
225	230	235
Leu Tyr Phe Gln Tyr Val Tyr Trp Leu His Met Leu Tyr Ala Ala Leu		
245	250	255
Gly Ala Ile Cys Phe Thr Leu Phe Leu Ala Tyr Asp Thr Gln Leu Val		
260	265	270
Leu Gly Asn Arg Lys His Thr Ile Ser Pro Glu Asp Tyr Ile Thr Gly		
275	280	285
Ala Leu Gln Ile Tyr Thr Asp Ile Ile Tyr Ile Phe Thr Phe Val Leu		
290	295	300
Gln Leu Met Gly Asp Arg Asn		
305	310	

【0052】

配列番号：3

配列の長さ：383

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル：No

起源：

生物名：ホモ＝サピエンス

細胞の種類：リンホーマ

セルライン：U937

クローン名：HP10493

配列

Met Ala Gly Ile Pro Gly Leu Leu Phe Leu Leu Phe Leu Leu Cys			
1	5	10	15
Ala Val Gly Gln Val Ser Pro Tyr Ser Ala Pro Trp Lys Pro Thr Trp			
20	25	30	
Pro Ala Tyr Arg Leu Pro Val Val Leu Pro Gln Ser Thr Leu Asn Leu			
35	40	45	
Ala Lys Pro Asp Phe Gly Ala Glu Ala Lys Leu Glu Val Ser Ser Ser			
50	55	60	
Cys Gly Pro Gln Cys His Lys Gly Thr Pro Leu Pro Thr Tyr Glu Glu			
65	70	75	80
Ala Lys Gln Tyr Leu Ser Tyr Glu Thr Leu Tyr Ala Asn Gly Ser Arg			
85	90	95	
Thr Glu Thr Gln Val Gly Ile Tyr Ile Leu Ser Ser Ser Gly Asp Gly			
100	105	110	
Ala Gln His Arg Asp Ser Gly Ser Ser Gly Lys Ser Arg Arg Lys Arg			
115	120	125	
Gln Ile Tyr Gly Tyr Asp Ser Arg Phe Ser Ile Phe Gly Lys Asp Phe			
130	135	140	
Leu Leu Asn Tyr Pro Phe Ser Thr Ser Val Lys Leu Ser Thr Gly Cys			
145	150	155	160
Thr Gly Thr Leu Val Ala Glu Lys His Val Leu Thr Ala Ala His Cys			
165	170	175	
Ile His Asp Gly Lys Thr Tyr Val Lys Gly Thr Gln Lys Leu Arg Val			
180	185	190	
Gly Phe Leu Lys Pro Lys Phe Lys Asp Gly Gly Arg Gly Ala Asn Asp			
195	200	205	

Ser Thr Ser Ala Met Pro Glu Gln Met Lys Phe Gln Trp Ile Arg Val
 210 215 220
 Lys Arg Thr His Val Pro Lys Gly Trp Ile Lys Gly Asn Ala Asn Asp
 225 230 235 240
 Ile Gly Met Asp Tyr Asp Tyr Ala Leu Leu Glu Leu Lys Lys Pro His
 245 250 255
 Lys Arg Lys Phe Met Lys Ile Gly Val Ser Pro Pro Ala Lys Gln Leu
 260 265 270
 Pro Gly Gly Arg Ile His Phe Ser Gly Tyr Asp Asn Asp Arg Pro Gly
 275 280 285
 Asn Leu Val Tyr Arg Phe Cys Asp Val Lys Asp Glu Thr Tyr Asp Leu
 290 295 300
 Leu Tyr Gln Gln Cys Asp Ala Gln Pro Gly Ala Ser Gly Ser Gly Val
 305 310 315 320
 Tyr Val Arg Met Trp Lys Arg Gln Gln Gln Lys Trp Glu Arg Lys Ile
 325 330 335
 Ile Gly Ile Phe Ser Gly His Gln Trp Val Asp Met Asn Gly Ser Pro
 340 345 350
 Gln Asp Phe Asn Val Ala Val Arg Ile Thr Pro Leu Lys Tyr Ala Gln
 355 360 365
 Ile Cys Tyr Trp Ile Lys Gly Asn Tyr Leu Asp Cys Arg Glu Gly
 370 375 380

【0053】

配列番号：4

配列の長さ：807

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名：H P O 1 2 0 7

配列

ATGGGCCAGA ACGACCTGAT GGGCACGGCC GAGGACTTCG CCGACCAGTT CCTCCGTGTC	60
ACAAAGCAGT ACCTGCCCA CGTGGCGCGC CTCTGTCTGA TCAGCACCTT CCTGGAGGAC	120
GGCATCCGTA TGTGGTTCCA GTGGAGCGAG CAGCGCGACT ACATCGACAC CACCTGGAAC	180
TGCGGCTACC TGCTGGCCTC GTCCCTCGTC TTCCCTCAACT TGCTGGGACA GCTGACTGGC	240
TGCGTCCCTGG TGTTGAGCAG GAACTTCGTG CAGTACGCCCT GCTTCGGGCT CTTTGGAAATC	300
ATAGCTCTGC AGACGATTGC CTACAGCATT TTATGGGACT TGAAGTTTT GATGAGGAAC	360
CTGGCCCTGG GAGGAGGCCT GTTGCTGCTC CTAGCAGAAT CCCGTTCTGA AGGGAAGAGC	420
ATGTTTGCAG GCGTCCCCAC CATGCGTGAG AGCTCCCCCA AACAGTACAT GCAGCTCGGA	480
GGCAGGGTCT TGCTGGTTCT GATGTTCATG ACCCTCCCTTC ACTTTGACGC CAGCTTCTTT	540
TCTATTGTCC AGAACATCGT GGGCACAGCT CTGATGATT TAGTGGCCAT TGGTTTTAAA	600
ACCAAGCTGG CTGCTTGAC TCTTGTGTG TGGCTCTTG CCATCAACGT ATATTCAAC	660
GCCTTCTGGA CCATTCCAGT CTACAAGCCC ATGCATGACT TCCTGAAATA CGACTTCTTC	720
CAGACCATGT CGGTGATTGG GGGCTTGCTC CTGGTGGTGG CCCTGGGCC C TGGGGGTGTC	780
TCCATGGATG AGAAGAAGAA GGAGTGG	807

【0054】

配列番号：5

配列の長さ：933

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名: H P 0 1 8 6 2

配列

ATGTCCAACC CCAGCGCCCC ACCACCATAT GAAGACCGCA ACCCCCTGTA CCCAGGCCCT	60
CTGCCCCCTG GGGGCTATGG GCAGCCATCT GTCCTGCCAG GAGGGTATCC TGCCTACCCCT	120
GGCTACCCGC AGCCTGGCTA CGGTACCCCT GCTGGCTACC CACAGCCCAT GCCCCCCACC	180
CACCCGATGC CCATGAACTA CGGCCAGGC CATGGCTATG ATGGGGAGGA GAGAGCGGTG	240
AGTGTAGCT TCGGGCCTGG AGAGTGGGAT GACCGGAAAG TGCGACACAC TTTTATCCGA	300
AAGGTTTACT CCATCATCTC CGTGCAGCTG CTCATCACTG TGGCCATCAT TGCTATCTTC	360
ACCTTGTGG AACCTGTCAG CGCCTTGTG AGGAGAAATG TGGCTGTCTA CTACGTGTCC	420
TATGCTGTCT TCGTTGTCAC CTACCTGATC CTTGCCTGCT GCCAGGGACC CAGACGCCGT	480
TTCCCATGGA ACATCATTCT GCTGACCCCTT TTTACTTTG CCATGGGCTT CATGACGGGC	540
ACCATTTCCA GTATGTACCA AACCAAAGCC GTCATCATTG CAATGATCAT CACTGCGGTG	600
GTATCCATT CAGTCACCAT CTTCTGCTT CAGACCAAGG TGGACTTCAC CTCGTGCACA	660
GGCCTCTTCT GTGTCTGGG AATTGTGCTC CTGGTGAUT GGATTGTAC TAGCATTGTG	720
CTCTACTTCC AATACGTTA CTGGCTCCAC ATGCTCTATG CTGCTCTGGG GGCCATTGT	780
TTCACCCCTGT TCCTGGCTTA CGACACACAG CTGGTCCTGG GGAACCGGAA GCACACCATC	840
AGCCCCGAGG ACTACATCAC TGGGCCCTG CAGATTACA CAGACATCAT CTACATCTTC	900
ACCTTGTGC TGCAGCTGAT GGGGGATCGC AAT	933

【0055】

配列番号: 6

配列の長さ: 1149

配列の型: 核酸

鎖の数: 二本鎖

トポロジー: 直鎖状

配列の種類: cDNA to mRNA

起源:

生物名: ホモ=サピエンス

細胞の種類: リンホーマ

セルライン: U937

クローン名: HP10493

配列

ATGGCAGGGA TTCCAGGGCT CCTCTTCCTT CTCTTCTTTC TGCTCTGTGC TGTTGGGCAA	60
GTGAGCCCTT ACAGTCCCCC CTGGAAACCC ACTTGGCCTG CATAACGCCT CCCTGTCGTC	120
TTGCCCCAGT CTACCCCTCAA TTTAGCCAAG CCAGACTTTG GAGCCGAAGC CAAATTAGAA	180
GTATCTTCTT CATGTGGACC CCAGTGTCAAT AAGGGAACTC CACTGCCAC TTACGAAGAG	240
GCCAAGCAAT ATCTGTCTTA TGAAACGCTC TATGCCAATG GCAGCCGCAC AGAGACGCAG	300
GTGGGCATCT ACATCCTCAG CAGTAGTGGAA GATGGGGCCC AACACCGAGA CTCAGGGTCT	360
TCAGGAAAGT CTCGAAGGAA GCGGCAGATT TATGGCTATG ACAGCAGGTT CAGCATTTC	420
GGGAAGGACT TCCTGCTCAA CTACCCCTTC TCAACATCAG TGAAGTTATC CACGGGCTGC	480
ACCGGCACCC TGGTGGCAGA GAAGCATGTC CTCACAGCTG CCCACTGCAT ACACGATGGA	540
AAAACCTATG TGAAAGGAAC CCAGAAGCTT CGAGTGGGCT TCCTAAAGCC CAAGTTAAA	600
GATGGTGGTC GAGGGGCCAA CGACTCCACT TCAGCCATGC CCGAGCAGAT GAAATTCAG	660
TGGATCCGGG TGAAACGCAC CCATGTGCCA AAGGGTTGGA TCAAGGGCAA TGCCAATGAC	720
ATCGGCATGG ATTATGATTA TGCCCTCCTG GAACTCAAAA AGCCCCACAA GAGAAAATT	780
ATGAAGATTG GGGTGAGCCC TCCTGCTAAG CAGCTGCCAG GGGGCAGAAT TCACTTCTCT	840
GGTTATGACA ATGACCGACC AGGCAATTG GTGTATCGCT TCTGTGACGT CAAAGACGAG	900
ACCTATGACT TGCTCTACCA GCAATGCGAT GCCCAGCCAG GGGCCAGCGG GTCTGGGTC	960
TATGTGAGGA TGTGGAAGAG ACAGCAGCAG AAGTGGGAGC GAAAAATTAT TGGCATTTC	1020
TCAGGGCACC AGTGGGTGGA CATGAATGGT TCCCCACAGG ATTTAACGT GGCTGTCAGA	1080
ATCACTCCTC TCAAATATGC CCAGATTGCA TATTGGATTA AAGGAAACTA CCTGGATTGT	1140
AGGGAGGGG	1149

配列番号: 7

配列の長さ: 2938

配列の型: 核酸

鎖の数: 二本鎖

トポロジー: 直鎖状

配列の種類: cDNA to mRNA

起源:

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名：H P 0 1 2 0 7

配列の特徴：

特徴を表す記号：C D S

存在位置：1 0 1 . . 9 1 0

特徴を決定した方法：E

配列

AAAAAGGGCA	CTTCCTGTGG	AGGCCGCAGC	GGGTGCGGGC	GCCGACGGGC	GAGAGCCAGC	60
GAGCGAGCGA	GCGAGCCGAG	CCGAGCCTCC	CGCCGTCGCC	ATG GGC CAG AAC GAC		115
				Met Gly Gln Asn Asp		
				1	5	
CTG ATG GGC ACG GCC GAG GAC TTC GCC GAC CAG TTC CTC CGT GTC ACA						163
Leu Met Gly Thr Ala Glu Asp Phe Ala Asp Gln Phe Leu Arg Val Thr						
10	15	20				
AAG CAG TAC CTG CCC CAC GTG GCG CGC CTC TGT CTG ATC AGC ACC TTC						211
Lys Gln Tyr Leu Pro His Val Ala Arg Leu Cys Leu Ile Ser Thr Phe						
25	30	35				
CTG GAG GAC GGC ATC CGT ATG TGG TTC CAG TGG AGC GAG CAG CGC GAC						259
Leu Glu Asp Gly Ile Arg Met Trp Phe Gln Trp Ser Glu Gln Arg Asp						
40	45	50				
TAC ATC GAC ACC ACC TGG AAC TGC GGC TAC CTG CTG GCC TCG TCC TTC						307
Tyr Ile Asp Thr Trp Asn Cys Gly Tyr Leu Leu Ala Ser Ser Phe						
55	60	65				
GTC TTC CTC AAC TTG CTG GGA CAG CTG ACT GGC TGC GTC CTG GTG TTG						355
Val Phe Leu Asn Leu Leu Gly Gln Leu Thr Gly Cys Val Leu Val Leu						
70	75	80	85			
AGC AGG AAC TTC GTG CAG TAC GCC TGC TTC GGG CTC TTT GGA ATC ATA						403
Ser Arg Asn Phe Val Gln Tyr Ala Cys Phe Gly Leu Phe Gly Ile Ile						

90	95	100	
GCT CTG CAG ACG ATT GCC TAC AGC ATT TTA TGG GAC TTG AAG TTT TTG			451
Ala Leu Gln Thr Ile Ala Tyr Ser Ile Leu Trp Asp Leu Lys Phe Leu			
105	110	115	
ATG AGG AAC CTG GCC CTG GGA GGA GGC CTG TTG CTG CTC CTA GCA GAA			499
Met Arg Asn Leu Ala Leu Gly Gly Gly Leu Leu Leu Leu Ala Glu			
120	125	130	
TCC CGT TCT GAA GGG AAG AGC ATG TTT GCG GGC GTC CCC ACC ATG CGT			547
Ser Arg Ser Glu Gly Lys Ser Met Phe Ala Gly Val Pro Thr Met Arg			
135	140	145	
GAG AGC TCC CCC AAA CAG TAC ATG CAG CTC GGA GGC AGG GTC TTG CTG			595
Glu Ser Ser Pro Lys Gln Tyr Met Gln Leu Gly Gly Arg Val Leu Leu			
150	155	160	165
GTT CTG ATG TTC ATG ACC CTC CTT CAC TTT GAC GCC AGC TTC TTT TCT			643
Val Leu Met Phe Met Thr Leu Leu His Phe Asp Ala Ser Phe Phe Ser			
170	175	180	
ATT GTC CAG AAC ATC GTG GGC ACA GCT CTG ATG ATT TTA GTG GCC ATT			691
Ile Val Gln Asn Ile Val Gly Thr Ala Leu Met Ile Leu Val Ala Ile			
185	190	195	
GGT TTT AAA ACC AAG CTG GCT GCT TTG ACT CTT GTT GTG TGG CTC TTT			739
Gly Phe Lys Thr Lys Leu Ala Ala Leu Thr Leu Val Val Trp Leu Phe			
200	205	210	
GCC ATC AAC GTA TAT TTC AAC GCC TTC TGG ACC ATT CCA GTC TAC AAG			787
Ala Ile Asn Val Tyr Phe Asn Ala Phe Trp Thr Ile Pro Val Tyr Lys			
215	220	225	
CCC ATG CAT GAC TTC CTG AAA TAC GAC TTC TTC CAG ACC ATG TCG GTG			835
Pro Met His Asp Phe Leu Lys Tyr Asp Phe Phe Gln Thr Met Ser Val			
230	235	240	245
ATT GGG GGC TTG CTC CTG GTG GTG GCC CTG GGC CCT GGG GGT GTC TCC			883

Ile Gly Gly Leu Leu Leu Val Val Ala Leu Gly Pro Gly Gly Val Ser

250

255

260

ATG GAT GAG AAG AAG AAG GAG TGG TAA CAGTCACAGA TCCCTACCTG 930

Met Asp Glu Lys Lys Lys Glu Trp

265

CCTGGCTAAG ACCCGTGGCC GTCAAGGACT GGTCGGGGT GGATTCAACA AACTGCCAG 990
 CTTTTATGTA TCCTCTTCCG TTCCCCTCCC TTGGTAAAGG CACAGATGTT TTGAGAACTT 1050
 TATTTGCAGA GACACCTGAG AATCGATGGC TCAGTCTGCT CTGGAGCCAC AGTCTGGCGT 1110
 CTGACCCCTTC AGTGCAGGCC AGCCTGGCAG CTGGAAGCCT CCCCCACGCC GAGGCTTTGG 1170
 AGTGAACAGC CCGCTTGGCT GTGGCATCTC AGTCCTATT TTGAGTTTT TTGTGGGGT 1230
 ACAGGAGGGG GCCTCAAGC TGTACTGTGA GCAGACGCAT TGGTATTATC ATTCAAAGCA 1290
 GTCTCCCTCT TATTGTAAG TTTACATT TTAGCGGAAAC TACTAAATT TAATGGGTGG 1350
 TTCAGCCAAA CCTCAAAACA GTTAATCTCC CTGGTTAAA ATCACACCAG TGGCTTTGAT 1410
 GTTGTTCCTG CCCCCCATTG TATTTATAG GAATACTGAA AACATTTAGG GACACCCAAA 1470
 GAATGATGCA GTATTAAGG GGTGGTAGAA GCTGCTGTT ATGATAAAAG TCATCGGTCA 1530
 GAAAATCAGC TTGGATTGGT GCCAAGTGT TTATTGGGTAA ACACCCCTGGG AGTTTAGTA 1590
 GCTTGAGGCA AGGTGGAGGG GCAAGAAGTC CTTGGGAAG CTGCTGGTCT GGGTGCTGCT 1650
 GCCCTCCAAG CTGGCAGTGG GAAGGGCTAG TGAGACCACA CAGGGTAGC CCCAGCAGCA 1710
 GCACCCCTGCA AGCCAGCCTG GCCAGCTGCT CAGACCAGCT TGCAGAGCCG CAGCCGCTGT 1770
 GGGCAGGGGG TGTGGCAGGA GCTCCCAGCA CTGGAGACCC ACGGACTCAA CCCAGTTACC 1830
 TCACATGGGG CCTTTCTGA GCAAGGTCTC GAAAGCGCAG GCCGCCCTGG CTGAGCAGCA 1890
 CCGCCCTTTC CCAGCTGCAC TCGCCCTGTG GACAGCCCCG ACACACCACT TTCTGAGGC 1950
 TGTGCTCAC TCAGATTGTC CGTTGCTAT GCCGAATGCA GCCAAAATTC CTTTTACAA 2010
 TTTGTGATGC CTTACCGATT TGATCTTAAT CCTGTATTAA AAGTTTCTA ACACTGCCTT 2070
 ATACTGTGTT TCTCTTTTG GGGGAGCTTA ACTGCTTGTG GCTCCCTGTC GTCTGCACCA 2130
 TAGTAAATGC CACAAGGGTA GTCGAACACC TCTCTGGCCC CTAGACCTAT CTGGGGACAG 2190
 GCTGGCTCAG CCTGTCTCCA GGGCTGCTGC GGGCCAGCCC CGAGCCTGCC TCCCTCTTGG 2250
 CCTCTCATCC ATTGGCTCTG CAGGGCAGGG GTGAGGCAGG TTTCTGCTCA TAAGTGCTTT 2310
 TGGAAGTCAC CTACCTTTT AACACAGCCG AACTAGTCCC AACCGCTTTG CAAATATTCC 2370

CCTGGTAGCC TACTTCCTTA	CCCCCGAATA TTGGTAAGAT CGATCAATGG	CTTCAGGACA	2430
TGGGTTCTCT TCTCCTGTGA	TCATTCAAGT GCTCACTGCA	TGAAGACTGG	2490
TGTTTCAACC TCACCCAGGGC	TGTCTCTTGG TCCACACCTC	GCTCCCTGTT AGTGCCGTAT	2550
GACAGCCCCC ATCAAATGAC	CTTGGCCAAG TCACGGTTTC	TCTGTGGTCA AGGTTGGTTG	2610
GCTGATTGGT GGAAAGTAGG	GTGGACCAAA GGAGGCCACG	TGAGCAGTCA GCACCAGTTC	2670
TGCACCAGCA GCGCCTCCGT	CCTAGTGGGT GTTCCTGTT	CTCCTGGCCC TGGGTGGGCT	2730
AGGGCCTGAT TCGGGAAGAT	GCCTTGAG GGAGGGGAGG	ATAAGTGGGA TCTACCAATT	2790
GATTCTGGCA AAACAATTTC	TAAGATTTT TTGCTTTATG	TGGAAACAG ATCTAAATCT	2850
CATTTATGC TGTATTTAT	ATCTTAGTTG TGTGAAACAG	CGTTTGATT TTTGGAAACA	2910
CATCAAAATA AATAATGGCG	TTTGTGT		2938

【0056】

配列番号：8

配列の長さ：2290

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名：HP01862

配列の特徴：

特徴を表す記号：CDS

存在位置：81..1016

特徴を決定した方法：E

配列

ACACTCCGAG GCCAGGAACG	CTCCGTCTGG AACGGCGCAG	GTCCCAGCAG	CTGGGGTTCC	60
CCCTCAGCCC GTGAGCAGCC	ATG TCC AAC CCC AGC	GCC CCA CCA CCA	TAT GAA	113
Met Ser Asn Pro Ser Ala Pro Pro Pro Tyr Glu				

1	5	10	
GAC CGC AAC CCC CTG TAC CCA GGC CCT CTG CCC CCT GGG GGC TAT GGG			161
Asp Arg Asn Pro Leu Tyr Pro Gly Pro Leu Pro Pro Gly Gly Tyr Gly			
15	20	25	
CAG CCA TCT GTC CTG CCA GGA GGG TAT CCT GCC TAC CCT GGC TAC CCG			209
Gln Pro Ser Val Leu Pro Gly Gly Tyr Pro Ala Tyr Pro Gly Tyr Pro			
30	35	40	
CAG CCT GGC TAC GGT CAC CCT GCT GGC TAC CCA CAG CCC ATG CCC CCC			257
Gln Pro Gly Tyr Gly His Pro Ala Gly Tyr Pro Gln Pro Met Pro Pro			
45	50	55	
ACC CAC CCG ATG CCC ATG AAC TAC GGC CCA GGC CAT GGC TAT GAT GGG			305
Thr His Pro Met Pro Met Asn Tyr Gly Pro Gly His Gly Tyr Asp Gly			
60	65	70	75
GAG GAG AGA GCG GTG AGT GAT AGC TTC GGG CCT GGA GAG TGG GAT GAC			353
Glu Glu Arg Ala Val Ser Asp Ser Phe Gly Pro Gly Glu Trp Asp Asp			
80	85	90	
CGG AAA GTG CGA CAC ACT TTT ATC CGA AAG GTT TAC TCC ATC ATC TCC			401
Arg Lys Val Arg His Thr Phe Ile Arg Lys Val Tyr Ser Ile Ile Ser			
95	100	105	
GTG CAG CTG CTC ATC ACT GTG GCC ATC ATT GCT ATC TTC ACC TTT GTG			449
Val Gln Leu Leu Ile Thr Val Ala Ile Ile Ala Ile Phe Thr Phe Val			
110	115	120	
GAA CCT GTC AGC GCC TTT GTG AGG AGA AAT GTG GCT GTC TAC TAC GTG			497
Glu Pro Val Ser Ala Phe Val Arg Arg Asn Val Ala Val Tyr Tyr Val			
125	130	135	
TCC TAT GCT GTC TTC GTT GTC ACC TAC CTG ATC CTT GCC TGC TGC CAG			545
Ser Tyr Ala Val Phe Val Val Thr Tyr Leu Ile Leu Ala Cys Cys Gln			
140	145	150	155
GGA CCC AGA CGC CGT TTC CCA TGG AAC ATC ATT CTG CTG ACC CTT TTT			593

Gly Pro Arg Arg Arg Phe Pro Trp Asn Ile Ile Leu Leu Thr Leu Phe
 160 165 170
 ACT TTT GCC ATG GGC TTC ATG ACG GGC ACC ATT TCC AGT ATG TAC CAA 641
 Thr Phe Ala Met Gly Phe Met Thr Gly Thr Ile Ser Ser Met Tyr Gln
 175 180 185
 ACC AAA GCC GTC ATC ATT GCA ATG ATC ATC ACT GCG GTG GTA TCC ATT 689
 Thr Lys Ala Val Ile Ile Ala Met Ile Ile Thr Ala Val Val Ser Ile
 190 195 200
 TCA GTC ACC ATC TTC TGC TTT CAG ACC AAG GTG GAC TTC ACC TCG TGC 737
 Ser Val Thr Ile Phe Cys Phe Gln Thr Lys Val Asp Phe Thr Ser Cys
 205 210 215
 ACA GGC CTC TTC TGT GTC CTG GGA ATT GTG CTC CTG GTG ACT GGG ATT 785
 Thr Gly Leu Phe Cys Val Leu Gly Ile Val Leu Leu Val Thr Gly Ile
 220 225 230 235
 GTC ACT AGC ATT GTG CTC TAC TTC CAA TAC GTT TAC TGG CTC CAC ATG 833
 Val Thr Ser Ile Val Leu Tyr Phe Gln Tyr Val Tyr Trp Leu His Met
 240 245 250
 CTC TAT GCT GCT CTG GGG GCC ATT TGT TTC ACC CTG TTC CTG GCT TAC 881
 Leu Tyr Ala Ala Leu Gly Ala Ile Cys Phe Thr Leu Phe Leu Ala Tyr
 255 260 265
 GAC ACA CAG CTG GTC CTG GGG AAC CGG AAG CAC ACC ATC AGC CCC GAG 929
 Asp Thr Gln Leu Val Leu Gly Asn Arg Lys His Thr Ile Ser Pro Glu
 270 275 280
 GAC TAC ATC ACT GGC GCC CTG CAG ATT TAC ACA GAC ATC ATC TAC ATC 977
 Asp Tyr Ile Thr Gly Ala Leu Gln Ile Tyr Thr Asp Ile Ile Tyr Ile
 285 290 295
 TTC ACC TTT GTG CTG CAG CTG ATG GGG GAT CGC AAT TAAGGAG 1020
 Phe Thr Phe Val Leu Gln Leu Met Gly Asp Arg Asn
 300 305 310

CAAGCCCCA	TTTCACCCG	ATCCTGGGCT	CTCCCTCCA	AGCTAGAGGG	CTGGGCCCTA	1080
TGACTGTGGT	CTGGGCTTTA	GGCCCCTTTC	CTTCCCCTTG	AGTAACATGC	CCAGTTCCCT	1140
TTCTGTCTG	GAGACAGGTG	GCCTCTCTGG	CTATGGATGT	GTGGGTACTT	GGTGGGGACG	1200
GAGGAGCTAG	GGACTAACTG	TTGCTCTTGG	TGGGCTTGGC	AGGGACTAGG	CTGAAGATGT	1260
GTCTTCTCCC	CGCCACCTAC	TGTATGACAC	CACATTCTTC	CTAACAGCTG	GGGTTGTGAG	1320
GAATATGAAA	AGAGCCTATT	CGATAGCTAG	AAGGGAATAT	GAAAGGTAGA	AGTGAACCTCA	1380
AGGTCAACGAG	GTTCCCCTCC	CACCTCTGTC	ACAGGCTTCT	TGACTACGTA	GTTGGAGCTA	1440
TTCTTCCCC	CAGCAAAGCC	AGAGAGCTTT	GTCCCCGGCC	TCCTGGACAC	ATAGGCCATT	1500
ATCCTGTATT	CCTTGCGCTT	GGCATTTTT	AGCTCAGGAA	GGTAGAAGAG	ATCTGTGCC	1560
ATGGGTCTCC	TTGCTTCAT	CCCTTCTTGT	TTCAGTGACA	TATGTATTGT	TTATCTGGGT	1620
TAGGGATGGG	GGACAGATAA	TAGAACGAGC	AAAGTAACCT	ATACAGGCCA	GCATGGAACA	1680
GCATCTCCCC	TGGGCTTGCT	CCTGGCTTGT	GACGCTATAA	GACAGAGCAG	GCCACATGTG	1740
GCCATCTGCT	CCCCATTCTT	GAAAGCTGCT	GGGGCCTCCT	TGCAGGCTTC	TGGATCTCTG	1800
GTCAGAGTGA	ACTCTTGCTT	CCTGTATTCA	GGCAGCTCAG	AGCAGAAAGT	AAGGGGCAGA	1860
GTCATACGTG	TGGCCAGGAA	GTAGCCAGGG	TGAAGAGAGA	CTCGGTGCGG	GCAGGGAGAA	1920
TGCCTGGGGG	TCCCTCACCT	GGCTAGGGAG	ATACCGAACG	CTACTGTGGT	ACTGAAGACT	1980
TCTGGGTTCT	TTCCCTCTGC	TAACCCAGGG	AGGGTCCTAA	GAGGAAGGTG	ACTTCTCTCT	2040
GTTTGTCTTA	AGTTGCACTG	GGGGATTCT	GACTTGAGGC	CCATCTCTCC	AGCCAGCCAC	2100
TGCCTTCTTT	GTAATATTAA	GTGCCTTGAG	CTGGAATGGG	GAAGGGGGAC	AAGGGTCAGT	2160
CTGTCGGGTG	GGGGCAGAAA	TCAAATCAGC	CCAAGGATAT	AGTTAGGATT	AATTACTTAA	2220
TAGAGAAATC	CTAACTATAT	CACACAAAGG	GATACAACTA	TAAATGTAAT	AAAATTTATG	2280
TCTAGAAGTT						2290

【0057】

配列番号：9

配列の長さ：3705

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA t o mRNA

起源：

生物名：ホモ＝サピエンス

セルライン：U 9 3 7

クローン名：H P 1 0 4 9 3

配列の特徴：

特徴を表す記号：C D S

存在位置：1 2 4 . . 1 2 7 5

特徴を決定した方法：E

配列

ACTCTCGGCT	GTGCCGGCGGG	GCAGGCATGG	GAGCCGGCGCG	CTCTCTCCCG	GCGCCCACAC	60
CTGTCTGAGC	GGCGCAGCGA	GCCGCGGCC	GGGCAGGCTG	CTCGGCGCGG	AACAGTGCTC	120
GGC ATG GCA GGG ATT CCA GGG CTC CTC TTC CTT CTC TTC TTT CTG CTC						168
Met Ala Gly Ile Pro Gly Leu Leu Phe Leu Leu Phe Phe Leu Leu						
1	5	10	15			
TGT GCT GTT GGG CAA GTG AGC CCT TAC AGT GCC CCC TGG AAA CCC ACT						216
Cys Ala Val Gly Gln Val Ser Pro Tyr Ser Ala Pro Trp Lys Pro Thr						
20	25	30				
TGG CCT GCA TAC CGC CTC CCT GTC GTC TTG CCC CAG TCT ACC CTC AAT						264
Trp Pro Ala Tyr Arg Leu Pro Val Val Leu Pro Gln Ser Thr Leu Asn						
35	40	45				
TTA GCC AAG CCA GAC TTT GGA GCC GAA GCC AAA TTA GAA GTA TCT TCT						312
Leu Ala Lys Pro Asp Phe Gly Ala Glu Ala Lys Leu Glu Val Ser Ser						
50	55	60				
TCA TGT GGA CCC CAG TGT CAT AAG GGA ACT CCA CTG CCC ACT TAC GAA						360
Ser Cys Gly Pro Gln Cys His Lys Gly Thr Pro Leu Pro Thr Tyr Glu						
65	70	75				
GAG GCC AAG CAA TAT CTG TCT TAT GAA ACG CTC TAT GCC AAT GGC AGC						408
Glu Ala Lys Gln Tyr Leu Ser Tyr Glu Thr Leu Tyr Ala Asn Gly Ser						
80	85	90	95			

CGC ACA GAG ACG CAG GTG GGC ATC TAC ATC CTC AGC AGT AGT GGA GAT 456
 Arg Thr Glu Thr Gln Val Gly Ile Tyr Ile Leu Ser Ser Ser Gly Asp
 100 105 110
 GGG GCC CAA CAC CGA GAC TCA GGG TCT TCA GGA AAG TCT CGA AGG AAG 504
 Gly Ala Gln His Arg Asp Ser Gly Ser Ser Gly Lys Ser Arg Arg Lys
 115 120 125
 CGG CAG ATT TAT GGC TAT GAC AGC AGG TTC AGC ATT TTT GGG AAG GAC 552
 Arg Gln Ile Tyr Gly Tyr Asp Ser Arg Phe Ser Ile Phe Gly Lys Asp
 130 135 140
 TTC CTG CTC AAC TAC CCT TTC TCA ACA TCA GTG AAG TTA TCC ACG GGC 600
 Phe Leu Leu Asn Tyr Pro Phe Ser Thr Ser Val Lys Leu Ser Thr Gly
 145 150 155
 TGC ACC GGC ACC CTG GTG GCA GAG AAG CAT GTC CTC ACA GCT GCC CAC 648
 Cys Thr Gly Thr Leu Val Ala Glu Lys His Val Leu Thr Ala Ala His
 160 165 170 175
 TGC ATA CAC GAT GGA AAA ACC TAT GTG AAA GGA ACC CAG AAG CTT CGA 696
 Cys Ile His Asp Gly Lys Thr Tyr Val Lys Gly Thr Gln Lys Leu Arg
 180 185 190
 GTG GGC TTC CTA AAG CCC AAG TTT AAA GAT GGT GGT CGA GGG GCC AAC 744
 Val Gly Phe Leu Lys Pro Lys Phe Lys Asp Gly Gly Arg Gly Ala Asn
 195 200 205
 GAC TCC ACT TCA GCC ATG CCC GAG CAG ATG AAA TTT CAG TGG ATC CGG 792
 Asp Ser Thr Ser Ala Met Pro Glu Gln Met Lys Phe Gln Trp Ile Arg
 210 215 220
 GTG AAA CGC ACC CAT GTG CCC AAG GGT TGG ATC AAG GGC AAT GCC AAT 840
 Val Lys Arg Thr His Val Pro Lys Gly Trp Ile Lys Gly Asn Ala Asn
 225 230 235
 GAC ATC GGC ATG GAT TAT GAT TAT GCC CTC CTG GAA CTC AAA AAG CCC 888
 Asp Ile Gly Met Asp Tyr Asp Tyr Ala Leu Leu Glu Leu Lys Lys Pro

240	245	250	255	
CAC AAG AGA AAA TTT ATG AAG ATT GGG GTG AGC CCT CCT GCT AAG CAG				936
His Lys Arg Lys Phe Met Lys Ile Gly Val Ser Pro Pro Ala Lys Gln				
260	265	270		
CTG CCA GGG GGC AGA ATT CAC TTC TCT GGT TAT GAC AAT GAC CGA CCA				984
Leu Pro Gly Gly Arg Ile His Phe Ser Gly Tyr Asp Asn Asp Arg Pro				
275	280	285		
GGC AAT TTG GTG TAT CGC TTC TGT GAC GTC AAA GAC GAG ACC TAT GAC				1032
Gly Asn Leu Val Tyr Arg Phe Cys Asp Val Lys Asp Glu Thr Tyr Asp				
290	295	300		
TTG CTC TAC CAG CAA TGC GAT GCC CAG CCA GGG GCC AGC GGG TCT GGG				1080
Leu Leu Tyr Gln Gln Cys Asp Ala Gln Pro Gly Ala Ser Gly Ser Gly				
305	310	315		
GTC TAT GTG AGG ATG TGG AAG AGA CAG CAG CAG AAG TGG GAG CGA AAA				1128
Val Tyr Val Arg Met Trp Lys Arg Gln Gln Gln Lys Trp Glu Arg Lys				
320	325	330	335	
ATT ATT GGC ATT TTT TCA GGG CAC CAG TGG GTG GAC ATG AAT GGT TCC				1176
Ile Ile Gly Ile Phe Ser Gly His Gln Trp Val Asp Met Asn Gly Ser				
340	345	350		
CCA CAG GAT TTC AAC GTG GCT GTC AGA ATC ACT CCT CTC AAA TAT GCC				1224
Pro Gln Asp Phe Asn Val Ala Val Arg Ile Thr Pro Leu Lys Tyr Ala				
355	360	365		
CAG ATT TGC TAT TGG ATT AAA GGA AAC TAC CTG GAT TGT AGG GAG GGG				1272
Gln Ile Cys Tyr Trp Ile Lys Gly Asn Tyr Leu Asp Cys Arg Glu Gly				
370	375	380		
TGACACAG TGTTCCCTCC TGGCAGCAAT TAAGGGTCTT CATGTTCTTA TTTTAGGAGA				1330
GGCCAAATTG TTTTTGTCA TTGGCGTGCA CACGTGTGTG TGTGTGTGTG TGTGTAAGGT				1390
GTCTTATAAT CTTTACCTA TTTCTTACAA TTGCAAGATG ACTGGCTTTA CTATTTGAAA				1450
ACTGGTTTGT GTATCATATC ATATATCATT TAAGCAGTTT GAAGGCATAC TTTTGCATAG				1510

AAATAAAAAA AATACTGATT TGGGGCAATG AGGAATATTT GACAATTAAG TTAATCTTCA	1570
CGTTTTGCA AACTTGATT TTTATTCAT CTGAACCTGT TTCAAAGATT TATATTAAAT	1630
ATTGGCATA CAAGAGATAT GAATTCTTAT ATGTGTGCAT GTGTGTTTC TTCTGAGATT	1690
CATCTTGGTG GTGGGTTTT TTGTTTTT AATTCACTGC CTGATCTTA ATGCTTCCAT	1750
AAGGCAGTGT TCCCATTAG GAACTTGAC AGCATTGTT AGGCAGAATA TTTGGATT	1810
GGAGGCATT GCATGGTAGT CTTGAACAG TAAAATGATG TGTTGACTAT ACTGATAACAC	1870
ATATTAAACT ATACCTTATA GTAAACCAGT ATCCCAAGCT GCTTTAGTT CCAAAAATAG	1930
TTCTTTCC AAAGGTTGTT GCTCTACTTT GTAGGAAGTC TTTGCATATG GCCCTCCAA	1990
CTTAAAGTC ATACCAGAGT GGCCAAGAGT GTTATCCCA ACCCTCCAT TTAACAGGAT	2050
TTCACTCACA TTTCTGGAAC TAGCTATTT TCAGAAGACA ATAATCAGGG CTTAATTAGA	2110
ACAGGCTGTA TTCCCTCCCA GCAAACAGTT GTGGCCACAC TAAAAACAAT CATAGCATT	2170
TACCCCTGGA TTATAGCACA TCTCATGTT TATCATTGG ATGGAGTAAT TTAAAATGAA	2230
TTAAATTCCA GAGAACAAATG GAAGCATTGC CTGGCAGATG TCACAACAGA ATAACCACCT	2290
GTTTGGAGCC TGGCACAGTC CTCCAGCCTG ATCAAAAATT ATTCTGCATA GTTTCAGTG	2350
TGCTTCTGG GAGCTATGTA CTTCTTCAAT TTGGAAACTT TTCTCTCTA TTTATAGTGA	2410
AAATACTTGG AAGTTACTTT AAGAAAACCA GTGTGGCCTT TTTCCCTCTA GCTTTAAAAG	2470
GGCCGTTTT GCTGGAATGC TCTAGGTTAT AGATAAACAA TTAGGTATAA TAGCAAAAAT	2530
GAAAATTGGA AGAATGCAAA ATGGATCAGA ATCATGCCTT CCAATAAAGG CCTTTACACA	2590
TGTTTATCA ATATGATTAT CAAATCACAG CATATACAGA AAAGACTTGG ACTTATTGTA	2650
TGTTTTATT TTATGGCTCT CGGCCTAAGC ACTTCTTCT AAATGTATCG GAGAAAAAAT	2710
CAAATGGACT ACAAGCACGT GTTGCTGTG CTTGCACCCC AGGTAAACCT GCATTGTAGC	2770
AATTGTAAG GATATTCAAG TGGAGCACTG TCACTTAGAC ATTCTCTGGG GGATTTCTG	2830
CTTGTCTTC TTGAGCTTT TGGAAGGATA ATTCTGATAA GGCACATCAAG AAACGTACAA	2890
CCACAGTGCT TTCTCAAAT CATATGAGAA ATACTATGCA TAGCAAGGAG ATGCAGAGCC	2950
GCCAGGAAAA TTCTGAGTTC CAGCACAATT TTCTTGGAA TCTAACAGGA ATCTAGCCTG	3010
AGGAAGAAGG GAGGTCTCCA TTTCTATGTC TGGTATTGG GGGTTTGTT TGTTTTGCT	3070
TTAGCTTGGT GAAAAAAAGT TCACTGAACA CCAAGACCAG AATGGATTAA TTTAAAAAAA	3130
TAGATGTTCC TTTGTGAAG CACCTTGATT CCTTGATT GATTTTGCA AAAGTTAGAC	3190
AATGGCACAA AGTCAAAATG AAATCAATGT TTAGTTACACA AGTAGATGTA ATTTACTAAA	3250

GAATGATACA CCCATATGCT ATATACAGCT TAACTCACAG AACTGTAAAA GAAAATTATA	3310
AAATAATTCA ACATGTCCAT CTTTTAGTG ATAATAAAAG AAAGCATGGT ATTAAACTAT	3370
CATAGAAGTA GACAGAAAAA GAAAAAAGGA CTCATGGCAT TATTAATATA ATTAGTGCTT	3430
TACATGTGTT AGTTATACAT ATTAGAAGCA TATTTGCCTA GTAAGGCTAG TAGAACCACA	3490
TTTCCCAAAG TGTGCTCCTT AAACACTCAT GCCTTATGAT TTTCTACCAA AAGTAAAAAG	3550
GGTTGTATTA AGTCAGAGGA AGATGCCTCT CCATTTCCC TCTCTTATC AGAGGTTCAC	3610
ATGCCTGTCT GCACATTAAA AGCTCTGGGA AGACCTGTTG TAAAGGGACA AGTTGAGGTT	3670
GTAAAATCTG CATTAAATA AACATCTTG ATCAC	3705

【0058】

【図面の簡単な説明】

【図1】 クローンH P 0 1 2 0 7がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

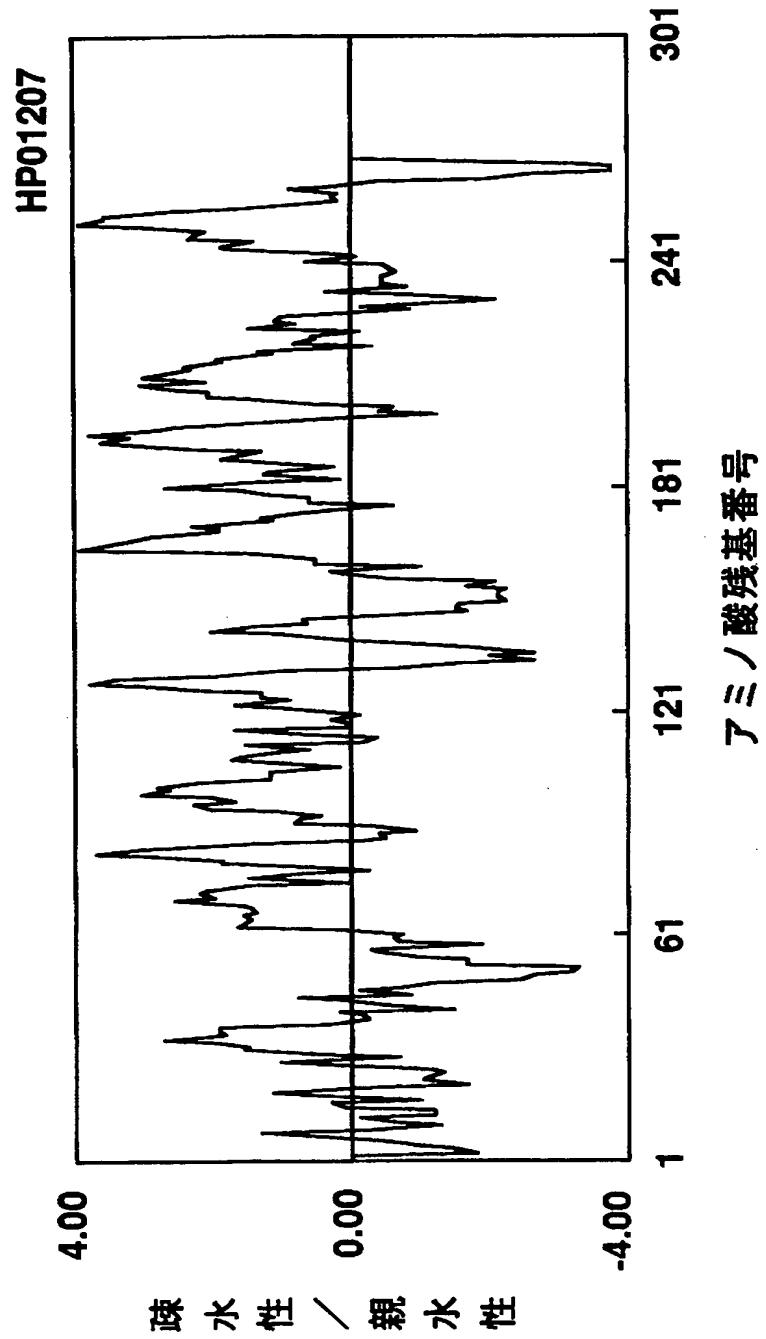
【図2】 クローンH P 0 1 8 6 2がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図3】 クローンH P 1 0 4 9 3がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

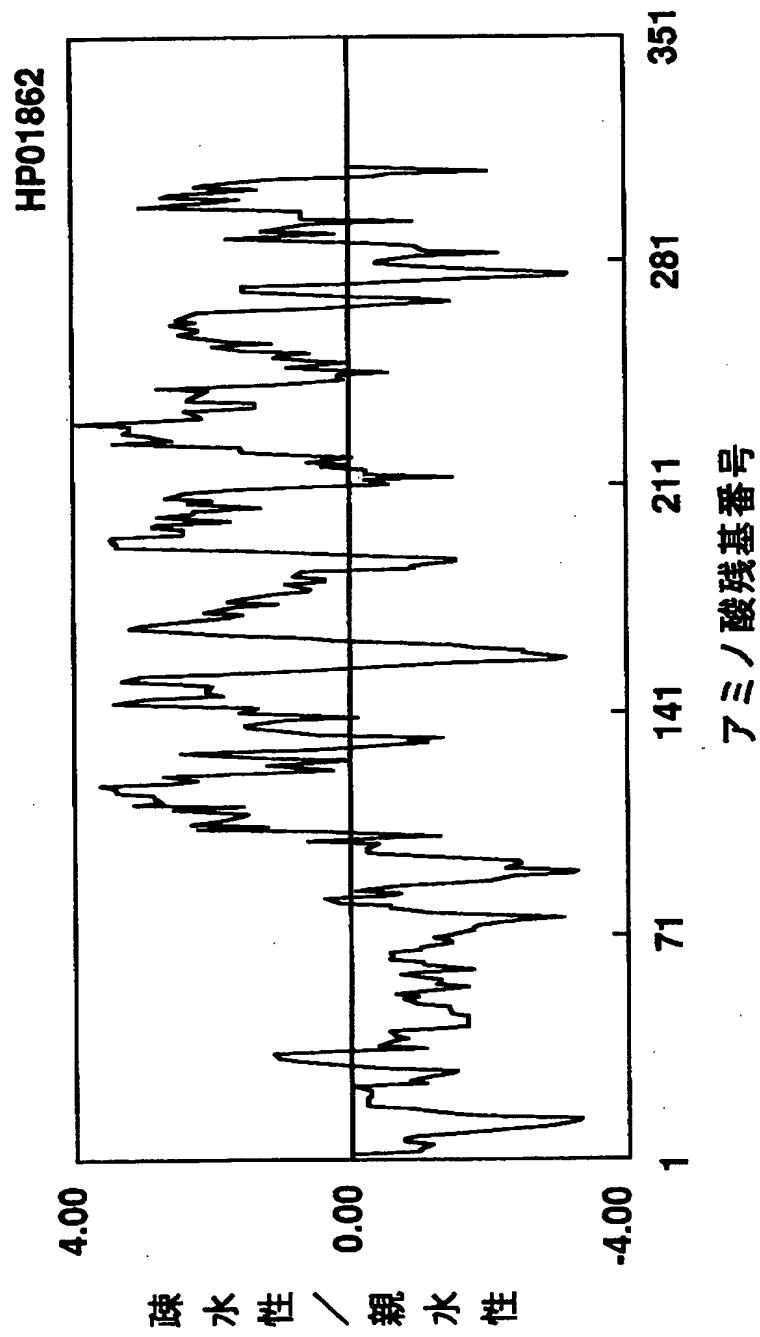
【書類名】

図面

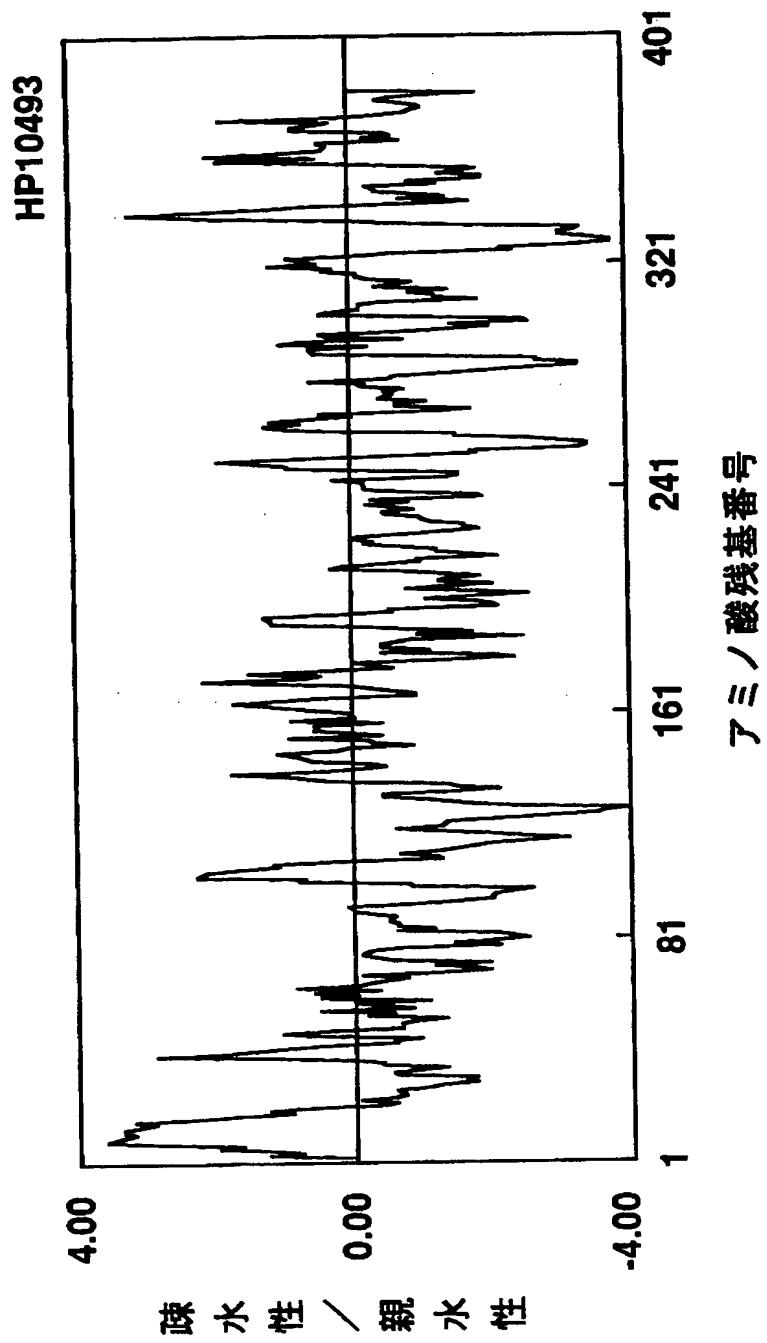
【図1】



【図2】



【図3】



【書類名】 要約書

【要約】

【課題】 膜貫通ドメインを有するヒト蛋白質、それをコードしている cDNA 、該 cDNA の発現ベクター、および該 cDNA を発現させた真核細胞を提供する。

【解決手段】 配列番号 1 から配列番号 3 で表されるアミノ酸配列のいずれかを含む蛋白質、該蛋白質をコードする DNA 、例えば配列番号 4 から配列番号 6 で表される塩基配列を含む cDNA 、該 cDNA の発現ベクター、および該 cDNA を発現させた真核細胞。膜貫通ドメインを有するヒト蛋白質をコードしている cDNA 、およびこのヒト cDNA の組換え体を発現させることにより該蛋白質ならびに該蛋白質を膜表面に有する真核細胞を提供することができる。

【選択図】 なし

【書類名】 職権訂正データ
【訂正書類】 特許願

〈認定情報・付加情報〉

【特許出願人】 申請人
【識別番号】 000173762
【住所又は居所】 神奈川県相模原市西大沼4丁目4番1号
【氏名又は名称】 財団法人相模中央化学研究所
【特許出願人】
【識別番号】 596134998
【住所又は居所】 東京都目黒区中町2丁目20番3号
【氏名又は名称】 株式会社プロテジーン

出願人履歴情報

識別番号 [000173762]

1. 変更年月日 1995年 4月14日

[変更理由] 住所変更

住 所 神奈川県相模原市西大沼4丁目4番1号

氏 名 財団法人相模中央化学研究所

出願人履歴情報

識別番号 [596134998]

1. 変更年月日 1996年 9月13日

[変更理由] 新規登録

住 所 東京都目黒区中町2丁目20番3号

氏 名 株式会社プロテジーン